

## DYNAMIC HOST CONFIGURATION PROTOCOL IN A STORAGE ENVIRONMENT

### BACKGROUND OF THE INVENTION

5    **1.    Field of the Invention**

The present invention relates generally to storage-area networks comprising diskless servers and computer data storage systems and, more specifically, to configuring a diskless server newly added to a storage-area network.

10   **2.    Description of the Related Art**

The common computing system paradigm for business enterprises has evolved from stand-alone computers, each operated by an individual user, to client-server systems in which a remote server computer provides the computing power to service client computers operated by users via a network. Client-server systems vary widely in the distribution of computing power, storage and other resources between the client and the server. For example, in some client-server systems, clients are termed "thin clients" because resources are concentrated in the server. In others, clients are powerful workstations having substantial storage and processing resources.

Another shift in the computing paradigm has been toward a model in which data storage is centralized. Individual workstations or other servers throughout an enterprise are networked to a central data storage subsystem that they use to a greater extent than their internal hard disks or even to the exclusion of internal disks. In some systems, the servers are completely diskless, with operating system software, application programs and data all stored in the central data storage subsystem. The network may be referred to as a storage-area network (SAN). The SAN model is particularly attractive for data centers and similar enterprises in which a substantial number of servers needs access to a central repository of data. Storage products for such systems are commercially available from a variety of sources, such as EMC Corporation of Hopkinton, Massachusetts.

## ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048

Systems are contemplated in which all of an enterprise's data storage needs are centralized, with the central data storage subsystem communicating remotely with diskless servers via a network protocol that supports communication over substantial distances, such as optical-media Fibre Channel. One advantage of diskless servers in a data center environment is that, if one fails, it can be quickly and easily replaced because there is no software to reinstall. Rather than having its operating system and other configuration software stored on a local disk, a diskless server dynamically obtains such software from the SAN during the boot process.

Installing new diskless servers in a data center environment, in which the enterprise may have many servers (hosts) running different operating systems, may tax the resources of a system administration department. If a server fails or otherwise needs to be replaced, or an additional server needs to be added, SAN ports and storage volumes must be assigned to the new host, and the operating system and application programs pertaining to the new host must then be installed on the SAN. This process often requires cooperation between the system administration department and a separate department in the enterprise that is charged with assigning ports and volumes and otherwise maintaining the central data storage devices. System administration personnel typically install operating systems and other software by loading from CD-ROMs or similar media or by using the so-called ghost image method by which software can be installed remotely via the network. Before beginning the installation process, such personnel must first identify the operating system that is to be installed and ensure that a copy is on hand. The overall process for adding a new host to a network can thus be time consuming and inefficient.

In contrast to the laborious process noted above for adding a new device to a SAN in a Fibre Channel or similar environment, adding devices to a network in an Internet Protocol (IP) network environment is increasingly easy with the advent of so-called "plug-and-play" concepts. In the Universal Plug and Play (UPnP) model, for example, promulgated by Microsoft Corporation, a newly connected device acts as a Dynamic Host

**ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048**

Configuration Protocol (DHCP) client and searches the network for a DHCP server. DHCP is a mechanism for providing device configuration information needed to access the Internet. If a DHCP server is available, the device obtains an IP address from it. The device then advertises its services to control points on the network. Similarly, when a control point is added to the network, the UPnP discovery protocol allows that control point to search for devices of interest on the network.

It would be desirable to be able to add a diskless server to a SAN with essentially the same ease as adding a plug-and-play device to an IP-based network. The present invention addresses this problem and others in the manner described below.

**SUMMARY OF THE INVENTION**

The present invention relates to automatic configuration of a diskless host computer upon adding it to storage-area network (SAN). Upon physically connecting the diskless host computer to the network, the host computer alerts the network and a control station of its presence. The control station uses a suitable identifier provided by the host, such as a Fibre Channel World Wide Name, to look up a configuration corresponding to a host of its type. Configurations corresponding to all types of hosts expected to be connected to the network are pre-stored in the SAN data storage system. Each configuration includes an operating system and can include any other software related to configuring or installing a new host. The control station then provides access to a storage device on which the operating system is stored so that the host can boot from it or access it for any other purposes.

It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048

**BRIEF DESCRIPTION OF THE DRAWINGS**

The accompanying drawings illustrate one or more embodiments of the invention and, together with the written description, serve to explain the principles of the invention. Wherever possible, the same reference numbers are used throughout the drawings to refer to the same or like elements of an embodiment, and wherein:

Figure 1 illustrates a SAN in which, upon physical connection of a diskless host, a Fibre Channel switch automatically couples the host to a device in the storage system on which configuration software for such a host is stored;

Figure 2 is a block diagram of a control station computer; and

Figure 3 is a flow diagram illustrating the method steps that occur upon connection of a diskless host to the SAN.

**DETAILED DESCRIPTION**

As illustrated in Fig. 1, a storage-area network (SAN) cloud 10 includes one or more data storage systems 12 and can communicate with one or more diskless hosts 14 via a Fibre Channel network 16. Data storage system 12 can be any suitable mass storage platform, such as the SYMMETRIX line of products from EMC Corporation of Hopkinton, Massachusetts. The SYMMETRIX is an enterprise data storage platform that includes the hardware and software necessary to serve the data storage needs of a business enterprise having many such hosts 14. The hardware includes an array of disk drives that typically have a collective storage capacity on the order of a terabyte or more. Nevertheless, the invention can be used in an environment having any suitable type of data storage system of any capacity. Note that hosts 14 are diskless; their persistent data storage needs are served entirely by SAN cloud 10 via Fibre Channel network 16. Because SAN cloud 10 is well-known in the art, its structure and operation are not described in further detail in this patent specification.

**ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048**

Fibre Channel network 16 is characterized by a Fibre Channel switch 18. A Fibre Channel switch is a well-known device commercially available from any of a number of manufacturers. Fibre Channel is a well-known network transport layer protocol that can be used with either optical (fiber) or electrical (copper wire) transmission media. For background information on the Fibre Channel standard, see ANSI X3.230-1994, "Fibre Channel - Physical and Signaling Interface" and related ANSI Fibre Channel standards. Fibre Channel can be used in various network topologies, but the topology created by using one or more switches 18 is known as a fabric. A well-known use of Fibre Channel is to allow several computers to share the same storage device. Therefore, this aspect of the illustrated system is not described in further detail in this patent specification.

Switch 18 has ports coupled (either directly or via one or more intermediate elements (not shown) of network 16) to SAN cloud 10, ports coupled to hosts 14, and one port coupled to a control station 20. Each port to SAN cloud 10 is associated with one of a number of individual logical storage devices 22, 24, 26, 28, 30, etc. As persons of skill in the art to which the invention pertains understand, switch 18 is of the crossbar type, such that any selected one of the ports to SAN cloud 10 can be connected to any selected one of the ports to hosts 14. Control station 20 provides control and port-selection signals to switch 18 that cause switch 18 to connect selected ports to each other, thereby connecting one of storage devices 22, 24, 26, 28 and 30 to one of hosts 14.

Note that the depiction of SAN cloud 10 as having devices 22, 24, 26, 28 and 30 separate from storage systems 12 is conceptual and for purposes of illustration; in actuality, devices 22, 24, 26, 28 and 30 are included in storage systems 12. Devices 22, 24, 26, 28 and 30 represent any suitable divisions of storage systems 12, whether physically separate disks or logically separate units or some combination thereof, that can be accessed independently of one another by hosts 14. Thus, the term "device" or "storage device" has the meaning commonly associated with it by persons of skill in the art. For reasons explained below, there is also a counterpart device 22', 24', 26', 28' and

30' to each of devices 22, 24, 26, 28 and 30. Counterpart devices 22', 24', 26', 28' and 30' are not necessarily physically adjacent or otherwise paired with devices 22, 24, 26, 28 and 30 in the manner shown; rather, they are conceptually shown that way in Fig. 1 for purposes of illustration.

5 Configuration data, including operating system software, is stored in each of devices 22, 24, 26, 28 and 30. Each stores configuration data different from that which the others store. The difference can be the operating system, with each storing a different operating system. For example, device 22 can store configuration data that includes a Sun Microsystems SOLARIS 7 operating system; device 24 can store configuration data  
10 that includes a SOLARIS 8 operating system; device 26 can store configuration data that includes a LINUX operating system, device 28 can store configuration that includes an AIX operating system, and so forth. Although SAN cloud 10 stores such configuration data for purposes of the present invention, its primary purpose is to store the data that relates to the enterprise's business activities. Those data are stored in still other devices  
15 32, 34, 36, 38 and 40.

The invention relates to automatically, i.e., without further human intervention, configuring host 14 when a system administrator or other person first physically connects it to network 16. Host 14 can be any suitable diskless server or other diskless computer whose storage needs can be provided by SAN 10 via network 16. Network 16 can be of  
20 any suitable type, but in the illustrated embodiment is Fibre Channel because Fibre Channel is commonly used in SAN environments. It is anticipated that other protocols, such as Internet Protocol (IP), will come into widespread commercial use in SAN environments as well. Because in the illustrated embodiment network 16 is Fibre Channel, host 14 has a conventional Fibre Channel host bus adapter (HBA) (not shown).  
25 It is through the HBA that host 14 can be physically connected to network 16. For reasons described below, host 14 also has a conventional IP network interface (not

**ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048**

shown) and is connected to control station 20 via an IP network 42. Host 14 can establish a connection to IP network 42 using DHCP in the conventional manner.

As illustrated in Fig. 2, control station 20 includes software 44 that reflects the method illustrated in Fig. 3, described below. Control station 20 can be any suitable type of general-purpose computer, such as a personal computer or a workstation. As such, it includes a main memory 46, a processor 48, suitable input/output (I/O) logic 50, and other hardware and software elements of the types conventionally included in such computers. (Such additional hardware and software elements are not shown in Fig. 2 for purposes of clarity.) Control station 20 also includes software for controlling and maintaining SAN cloud 10 but is not shown for purposes of clarity. In addition, control station 20 includes a Fibre Channel HBA 52 for communicating with switch 18 and an IP network interface 54 for communicating with host 14. I/O logic 50 communicates with a database or similar data store operating under the well-known lightweight directory access protocol (LDAP) 56. Although illustrated as a separate device for purposes of clarity, LDAP 56 can alternatively be integrated with control station 20 in any suitable manner. Control station 20 also includes other I/O elements of the types commonly included in computers, such as a keyboard, monitor and mouse but are not shown for purposes of clarity.

Although software 44 can be modularized or otherwise structured in any suitable manner using any suitable software technology and written in any suitable programming language, its constituent elements can be conceptually understood to include at least a switch control element 58 for controlling switch 18, a LDAP lookup element 60 for controlling LDAP 56, and a storage system interface element 62. Although software elements 58, 60 and 62 are conceptually illustrated as residing in memory 46 for execution by processor 48, as persons of skill in the art understand, all software elements may not all reside in memory 46 simultaneously or in their entirety. Furthermore, as such persons understand, there are other software elements of control station 20 that are

not shown for purposes of clarity, such as a suitable operating system, IP stack and other system software, but which would reside in memory 46 at some time at least in part. Also, control station 20 is not diskless, and software elements can be stored on a suitable disk storage device and retrieved on an as-needed basis in the conventional manner. This device can be the same as that of LDAP 56 or a different device. Note that the present invention can be embodied not only as a method and system but also as a computer program product that comprises a disk or other medium on which software 44 is stored or otherwise carried.

With regard to Fig. 3, at step 64 a system administrator or other person physically connects host 14 to Fibre Channel network 16, i.e., "plugs it in." In response to being plugged in to network 16, at step 66 the HBA of host 14 attempts to perform a Fibre Channel login process. The Fibre Channel login process is well-known in the art and therefore is not described in detail in this patent specification. Nevertheless, the following is a brief synopsis.

Attempting a login upon initial connection to a network is an inherent feature of all Fibre Channel HBAs because it is specified in the Fibre Channel ANSI standard. Fibre Channel uses a three byte address identifier as a network address, which is dynamically assigned during the login process. In the login process, the node port transmits a Fabric Login (FLOGI) frame to the well-known fabric address 'FFFFFF'. The fabric returns an Accept (ACC) frame back to the node port. As part of the login process, the host transmits its world wide name (WWN) to the fabric address. Every Fibre Channel HBA has a WWN, which is a 64-bit number assigned by the HBA manufacturer that uniquely defines the device. In other words, no two HBAs or other two Fibre Channel devices have the same WWN. A portion of the 64-bit WWN identifies the manufacturer.

Although host 14 itself has no operating system loaded at this point in the process, its HBA has an internal processor and read-only memory that can effect the login process independently of host 14. As part of the process of logging into switch 18, which in the



illustrated embodiment of the invention constitutes the fabric, the HBA of host 14 transmits its WWN to switch 18. In conformance with the Fibre Channel standard, switch 18 sends a message to all other ports to the effect that a device is attempting to log in. The message also specifies the WWN of the device and the port into which it is attempting to log in.

At step 68 control station 20 receives the WWN from switch 18. The WWN is used in the present invention to identify an appropriate operating system and other configuration data for host 14. HBAs designed for use in a computer having one type of operating system cannot be installed in computers having a different operating system. For example, one HBA product may be usable only in a UNIX environment and another may be usable only in a WINDOWS NT environment. Stored in LDAP 56 is a correspondence between the WWN of each HBA likely to be used in the system and configuration data. At step 70 control station 20 uses the WWN as a lookup key in LDAP 56 and finds the corresponding configuration data. Thus, for example, control station 20 may determine in response to a certain WWN it received from host 14 that host 14 is of a type that requires a SOLARIS 7 operating system and other configuration data. The manufacturer identification portion of the WWN may be all that is needed to identify a corresponding operating system because in many cases all of a manufacturer's computers operate under the same type of operating system. Although LDAP 56 is used to perform the lookup in the illustrated embodiment of the invention, in other embodiments other suitable lookup or database mechanisms can be used.

Alternatively to steps 68 and 70, wherein control station 20 receives the WWN from switch 18 and uses it to look up corresponding configuration data, control station 20 can receive directly from host 14 via IP network 42 an indication of the proper configuration. In other words, host 14 can send a message directly to control station 20 to inform control station 20 of the proper operating system and other configuration data. In such an embodiment of the invention, some minimal operating software, such as the

Microsoft Disk Operating System (MS-DOS or PC-DOS) and an IP stack, can be loaded from a floppy disk or CD-ROM (not shown) into host 14 before plugging host 14 into network 16 at step 64. The operating software is sufficient to enable host 14 to perform the initial task of transmitting a configuration message to control station 20. When host 14 is ultimately provided access to the proper operating system, as described below, the floppy disk or CD-ROM can be removed and host 14 re-booted.

In a still further alternative embodiment, which is somewhat of a hybrid between the two described above, host 14 can send a message to control station 20 via IP network 42 to inform control station 20 that host 14 is attempting to log in to switch 18. In response, control station 20 can receive the WWN from switch 18. Thus, in such an embodiment of the invention, IP network 42 is used to notify control station 20 of the login attempt but not to identify a configuration. Control station 20 obtains the configuration information in the same manner described above with regard to steps 68 and 70.

Note that each configuration, including an operating system and any other configuration data, is stored in a corresponding one of devices 22, 24, 26, 28 and 30. At step 72 control station 20 copies the configuration identified at step 70 from the corresponding one of devices 22, 24, 26, 28 and 30 to its counterpart device 22', 24', 26', 28' or 30'. Although not essential, this step is performed to preserve the integrity of the original copy of the configuration data. In other words, the configurations on devices 22, 24, 26, 28 and 30 are preserved as so-called "gold" copies, and the copies on devices 22', 24', 26', 28' and 30' become working copies. The working copies can then be altered to suit host requirements, such as name and IP address.

At step 74 control station 20 sends a message to switch 18 to establish a connection between two ports. One of the ports identified in the message is that which is coupled to the one of devices 22', 24', 26', 28' and 30' on which the identified configuration data are stored. The other port identified in the message is that which is

**ATTORNEY DOCKET NO.: 05125.0001U1/EMC 99-048**

coupled to the logged-in host 14 that requires the configuration data. Switch 18 thus establishes a path between host 14 and the one of devices 22', 24', 26', 28' and 30' on which its configuration data are stored. With the path established in this manner, host 14 has access to the operating system and other configuration data. At step 76 host 14 can boot from this copy of the operating system and otherwise configure itself in preparation for normal operation.

A system administrator need only plug host 14 into network 16, and host 14 will be ready to use as soon as the above-described steps have occurred. If host 14 requires replacement, or if an additional host 14 is to be added to the SAN, no time is spent by the system administrator installing software or otherwise configuring the new host 14.

It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.